Design Guidelines for Prompt Engineering Text-to-Image Generative Models

Abstract

Text-to-image generative models are a new and powerful way to generate visual artwork. However, the open-ended nature of text as interaction is double-edged; while users can input anything and have access to an infinite range of generations, they also must engage in brute-force trial and error with the text prompt when the result quality is poor. We conduct a study exploring what prompt keywords and model hyperparameters can help produce coherent outputs. In particular, we study prompts structured to include subject and style keywords and investigate success and failure modes of these prompts. Our evaluation of 5493 generations over the course of five experiments spans 51 abstract and concrete subjects as well as 51 abstract and figurative styles. From this evaluation, we present design guidelines that can help people produce better outcomes from text-to-image generative models.